



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 142220

To: Nita M Minnifield  
Location: rem/3c01/3c18  
Art Unit: 1645  
Monday, January 24, 2005

Case Serial Number: 09/837344

From: Beverly Shears  
Location: Remsen Bldg.  
RM 1A54  
Phone: 571-272-2528

beverly.shears@uspto.gov

### Search Notes

*Reviewed  
2-1-05  
mm.*



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 20:50:38 ; Search time 533 Seconds

(without alignments)  
14595.951 Million cell updates/sec

Title: US-09-837-344-41

Sequence: 1482  
1 CAAGACACACAAAGCATCT.....AATATTATTGAAACTATA 1482

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_23Sep04:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2000s:\*  
5: geneseqn2001as:\*  
6: geneseqn2001bs:\*  
7: geneseqn2002as:\*  
8: geneseqn2002bs:\*  
9: geneseqn2003as:\*  
10: geneseqn2003bs:\*  
11: geneseqn2003as:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1482	100.0	1496	2	AAQ28119
2	950	64.1	950	2	AAQ28115
3	853.2	57.6	1320	2	AAQ80916
4	586.4	39.6	1566	2	AAQ05140
5	503.4	34.0	1371	12	ADO21941
6	493.6	33.3	1374	12	ADO21919
7	472.8	31.9	988	2	AAQ28117
8	233.4	15.7	1300	12	ADP85917
9	225.8	15.2	3399	2	AAT05868
10	183.2	12.4	210	1	AAN08089
11	179.4	12.1	1686	2	AAQ80916
12	176.6	11.9	3579	3	AAQ70099
13	170.6	11.5	1998	3	AAA70212
14	151.2	10.2	6292	4	AAQ46735
15	149.2	10.1	6767	4	AAQ46735
16	146	9.9	3662	4	AAQ46735
17	139.4	9.4	9539	6	ABK28180
18	139.4	9.3	6668	6	ABU33697
19	137.8	9.2	6644	6	AAQ33181
20	136.4	9.2	7372	2	AAQ33182
21	136.4	9.2	7372	2	AAQ33182

22	136.4	9.2	7797	2	AAQ33180	AAQ33180 Cowpox vi
23	136.4	9.2	7996	2	AAQ33184	AAQ33184 Base sequ
24	132.8	9.0	778	6	ABQ15588	ABQ15588 Oligonuc
25	132.8	9.0	778	6	ABQ15589	ABQ15589 Oligonuc
26	131.4	8.9	7442	4	AAQ46686	AAQ46686 Tumour su
27	129.8	8.8	1762	6	ABQ20939	ABQ20939 Oligonuc
28	129.8	8.8	1762	6	ABQ20938	ABQ20938 Oligonuc
29	129.2	8.7	1995	6	ABQ20942	ABQ20942 Oligonuc
30	129.2	8.7	1995	6	ABQ20943	ABQ20943 Oligonuc
31	129	8.7	1200	6	ABQ39210	ABQ39210 Oligonuc
32	129	8.7	1200	6	ABQ39211	ABQ39211 Oligonuc
33	129	8.7	7597	6	ABQ33013	ABQ33013 Human imm
34	129	8.7	14006	6	ABQ33958	ABQ33958 Human imm
35	127.4	8.6	3683	8	ABQ10199	ABQ10199 Haematopo
36	127	8.6	3683	8	ABQ10053	ABQ10053 Haematopo
37	123.4	8.3	3037	6	ABQ54066	ABQ54066 Oligonuc
38	123.4	8.3	3037	6	ABQ54067	ABQ54067 Oligonuc
39	123	8.3	12237	6	ABQ34358	ABQ34358 Human imm
40	120.4	8.1	1533	8	ACQ28672	ACQ28672 Prokaryot
41	120.4	8.1	6419	6	ABQ32267	ABQ32267 Human imm
42	118.6	8.0	975	6	ABQ29508	ABQ29508 Oligonuc
43	118.6	8.0	975	6	ABQ29509	ABQ29509 Oligonuc
44	118.6	8.0	7758	6	ABQ33103	ABQ33103 Human imm
45	117.8	7.9	6898	6	ABQ80222	ABQ80222 Human che

## ALIGNMENTS

RESULT 1	AAQ28119	standard; DNA; 1496 BP.
ID	AAQ28119	
AC	AAQ28119;	
XX		
DT	25-MAR-2003	(revised)
XX		
DT	08-FEB-1993	(first entry)
XX		
DE	P.falciparum LSA gene 3' region.	
XX		
KM	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;	
XX		
KM	paludism; liver stage-specific antigen; ss.	
XX		
OS	Plasmodium falciparum.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1482
FT		/*tag= C
FT		/product= "LSA_C-terminal_region"
FT	repeat_region	37..639
FT		/*tag= a
FT	repeat_unit	37..87
FT		/*tag= b
FT		/rpt_type= TANDEM
XX		
PD	MO9213884-A1.	
XX		
PD	20-AUG-1992.	
XX		
PF	05-FEB-1992;	92MO-FR000104.
XX		
PR	05-FEB-1991;	91FR-00001286.
XX		
PA	(INSP ) INST PASTEUR.	
XX		
PI	Guerinmarchand C, Druilhe P;	
XX		
DR	WPI. 1992-299885/36.	
XX		
DR	P-PSDB; AAK26944.	
XX		
PT	Polypeptide(s) derived from liver stage of PLASMODIUM FALCIPARUM - for	
XX		
PT	vaccination against, treatment of and diagnosis of malaria.	
XX		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 23:08:59 ; Search time 101 Seconds  
(without alignments)  
10429.595 Million cell updates/sec

Title: US-09-837-344-41

Sequence: 1 CAGAGAACACAAAGCATCT.....AATTTTATGAACATAA 1482

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1482	100.0	1482	3	US-08-098-327E-41
2	1482	100.0	1482	3	US-08-462-625-41
3	1480.4	99.9	1482	3	US-08-098-327E-45
4	1480.4	99.9	1482	3	US-08-462-625-45
5	1446	97.6	1493	3	US-08-098-327E-38
6	1446	97.6	1493	3	US-08-462-625-38
7	950	64.1	950	3	US-08-098-327E-32
8	950	64.1	950	3	US-08-462-625-32
9	853.2	57.6	1320	1	US-08-257-073-15
10	472.8	31.9	988	3	US-08-098-327E-34
11	472.8	31.9	988	3	US-08-462-625-34
12	470.8	31.8	954	3	US-08-098-327E-37
13	470.8	31.8	954	3	US-08-462-625-37
14	116.2	7.8	3095	6	5231168-1
15	1837	7.5	1837	4	US-09-270-767-1308
16	111.6	7.5	1837	4	US-09-270-767-16550
17	110.2	7.4	7218	1	US-08-222-463-14
18	101.2	6.8	1956	4	US-08-559-8968-1
19	101.2	6.8	1956	4	US-09-351-794A-1
20	93.6	6.3	5361	3	US-08-973-462-2
21	93.6	6.3	5361	3	US-08-973-462-2
22	89.6	6.0	1664976	4	US-08-916-421B-1
23	89.6	6.0	1664976	4	US-09-692-570-1
24	88.4	6.0	5340	4	US-09-627-122-21
25	86.8	5.9	699	4	US-09-248-796A-9722
26	86.8	5.9	4766	5	PCT-US93-07261-10
27	83.2	5.6	1669	3	US-09-461-697-184

28	82.8	5.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
29	81.6	5.5	5455	4	US-10-204-708-33	Sequence 33, Appl
30	81.4	5.5	2223	1	US-08-257-073-4	Sequence 4, Appl
31	81.2	5.5	696	3	US-09-461-697-193	Sequence 193, App
32	81.2	5.5	699	3	US-09-461-697-191	Sequence 191, App
33	81.2	5.5	717	3	US-09-461-697-189	Sequence 189, App
34	81.2	5.5	774	3	US-09-461-697-187	Sequence 187, App
35	81.2	5.5	819	3	US-09-461-697-185	Sequence 185, App
36	80.2	5.4	6669	4	US-10-204-708-6	Sequence 6, Appl
37	79.6	5.4	53332	4	US-10-224-562-3	Sequence 3, Appl
38	79.6	5.4	8961	4	US-10-204-708-80	Sequence 80, Appl
39	78.4	5.3	8961	4	US-09-640-173-53	Sequence 53, Appl
40	78.2	5.3	10640	4	US-09-713-550-53	Sequence 53, Appl
41	76.8	5.2	396	4	US-09-825-294-53	Sequence 53, Appl
42	76.8	5.2	396	4	US-09-970-966-53	Sequence 53, Appl
43	76.8	5.2	396	4	US-09-970-966-53	Sequence 53, Appl
44	76.8	5.2	396	4	US-09-970-966-53	Sequence 53, Appl
45	76.8	5.2	705	4	US-09-270-767-5061	Sequence 5061, Ap

## ALIGNMENTS

RESULT 1  
US-08-098-327E-41  
Sequence 41, Application US/08098327E  
Patent No. 6270771  
GENERAL INFORMATION:  
APPLICANT: GUERIN-MARCHAND, Claudine  
APPLICANT: DRUILHE, Pierre  
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE  
TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Burns, Doane, Swecker & Machis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22113-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/098,327E  
FILING DATE: 24-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 91 01286  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 010830-045  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1482 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1482  
PUBLICATION INFORMATION:  
DOCUMENT NUMBER: WO 92/13884  
PUBLICATION DATE: 20-AUG-1992

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2005, 00:49:49 ; Search time 611 Seconds  
(without alignments)  
13936.836 Million cell updates/sec

Title: US-09-837-344-41

Sequence: 1 CAAGAACACCAAGCATCT.....AATATTATTGAACTATA 1482

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA:\*

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*

10: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*

14: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*

15: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*

16: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*

17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*

18: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*

19: /cgn2\_6/prodata/1/pubpna/US11\_NEW\_PUB.seq:\*

20: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*

21: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	1482	100.0	1482	9	US-09-837-344-41
2	1482	100.0	1482	10	US-09-900-963-41
3	1480.4	99.9	1482	9	US-09-837-344-45
4	1480.4	99.9	1482	10	US-09-900-963-45
5	1446	97.6	1493	9	US-09-837-344-38
6	1446	97.6	1493	10	US-09-900-963-38
7	950	64.1	950	9	US-09-837-344-32
8	950	64.1	950	10	US-09-900-963-32
9	472.8	31.9	988	9	US-09-837-344-34
10	472.8	31.9	988	10	US-09-900-963-34
11	470.8	31.8	954	9	US-09-837-344-37
12	470.8	31.8	954	10	US-09-900-963-37

13	233.4	15.7	1300	17	US-10-668-749A-1	Sequence 1, Appl
14	153.2	10.3	1297	18	US-10-425-115-67697	Sequence 67697, A
15	151.2	10.2	6292	16	US-10-221-714A-461	Sequence 461, App
16	149.2	10.1	6767	16	US-10-221-714A-330	Sequence 330, App
17	145.8	9.8	1243	18	US-10-425-115-172717	Sequence 172717, A
18	145	9.6	1204	17	US-10-437-963-77858	Sequence 77858, A
19	142.4	9.6	1081	18	US-10-425-115-16756	Sequence 16756, A
20	139.4	9.4	9539	14	US-10-239-676-52	Sequence 52, Appl
21	139.4	9.4	9539	15	US-10-240-453-54	Sequence 54, Appl
22	137.8	9.3	6668	15	US-10-311-455-1670	Sequence 1670, Ap
23	132.8	9.0	778	18	US-10-363-345A-2179	Sequence 2179, Ap
24	132.8	9.0	778	18	US-10-363-345A-2180	Sequence 2180, Ap
25	131.4	8.9	7442	16	US-10-221-714A-409	Sequence 409, App
26	129.8	8.8	521	18	US-10-425-115-36710	Sequence 36710, A
27	129.8	8.8	1762	18	US-10-363-345A-7529	Sequence 7529, Ap
28	129.8	8.8	1762	18	US-10-363-345A-7530	Sequence 7530, Ap
29	129.2	8.7	1595	18	US-10-363-345A-7533	Sequence 7533, Ap
30	129.2	8.7	1595	18	US-10-363-345A-7534	Sequence 7534, Ap
31	129	8.7	1200	18	US-10-363-345A-25801	Sequence 25801, A
32	129	8.7	1200	18	US-10-363-345A-25802	Sequence 25802, A
33	129	8.7	7597	15	US-10-311-455-1986	Sequence 986, App
34	129	8.7	14006	15	US-10-311-455-1931	Sequence 1931, App
35	128.6	8.7	3673778	15	US-10-312-841-1	Sequence 1, Appl
36	128.6	8.7	3673778	15	US-10-312-841-2	Sequence 2, Appl
37	127.4	8.6	3583	18	US-10-473-126-339	Sequence 339, App
38	127	8.6	1130	18	US-10-425-115-132399	Sequence 132399, A
39	127	8.6	3583	18	US-10-473-126-193	Sequence 193, App
40	126.2	8.5	880	18	US-10-425-115-53133	Sequence 53133, A
41	124.6	8.4	986	18	US-10-425-115-174478	Sequence 174478, A
42	123.8	8.4	1062	18	US-10-425-115-120013	Sequence 120013, A
43	123.4	8.3	3037	18	US-10-363-345A-40657	Sequence 40657, A
44	123.4	8.3	3037	18	US-10-363-345A-40658	Sequence 40658, A
45	123	8.3	12237	15	US-10-311-455-2231	Sequence 2231, App

#### ALIGNMENTS

RESULT 1  
US-09-837-344-41  
Sequence 41, Application US/09837344  
Patent No. US2002004182A1  
GENERAL INFORMATION:  
APPLICANT: GUERIN-MARCHAND, Claudine  
TITLE OF INVENTION: DRUGS, Pierre  
HEPATITIS STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE OF STIMULATING THE T LYMPHOCYTES  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/837,344  
FILING DATE: 19-Apr-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/462,625  
FILING DATE: <Unknown>  
APPLICATION NUMBER: FR 91 01286  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 23:02:29 ; Search time 3390 Seconds

(without alignments)  
15930.290 Million cell updates/sec

Title: US-09-837-344-41  
Sequence: 1482  
Sequence: 1 CAGGACACAGAACGATCT.....AATATTATGAACTATA 1482

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	291.4	19.7	317	9	T09652 0224m3 gmbp
2	248.4	16.8	1519	9	AG386893 Mus muscu
3	247.2	16.7	1453	9	AJ591978 Arabidops
4	241	16.3	1727	9	CU11085 ISB1-59J1
5	240	16.2	1491	9	CG753221 P048-2-A0
6	237.8	16.0	1434	9	AJ592058 Arabidops
7	237.8	15.9	1808	9	CU078963 CH216-154
8	235	15.9	1654	9	CU118709 ISB1-7211
9	234.2	15.8	1797	9	CU066150 CH216-108
10	234	15.8	1829	9	CU109409 ISB1-51H2
11	233.8	15.8	1736	9	CU066373 CH216-108
12	233.6	15.8	1407	9	AJ592026 Arabidops
13	232.8	15.7	1512	9	CU113943 ISB1-59E2
14	232.8	15.7	1531	9	CG748014 P041-4-B0
15	232.6	15.7	1811	9	CG753732 P048-4-G0
16	231.6	15.6	1917	9	CU081899 CH216-164
17	231.4	15.6	1650	9	CU041163 CH216-51P
18	231.2	15.6	1762	9	CU082587 CH216-168
19	230.4	15.5	1824	9	CU081234 CH216-161
20	230.2	15.5	1727	9	CU118815 ISB1-7205
21	230	15.5	1632	9	CU082569 CH216-165
22	229.8	15.5	1616	9	CU081995 CH216-152
23	229.6	15.5	1630	9	CU078686 CH216-152
24	229.2	15.5	1457	9	CU082658 CH216-169

C	25	228.8	15.4	1566	9	CG757757 P053-1-D0
C	26	228.6	15.4	1434	8	CU187638 CH261-98P
C	27	228.4	15.4	1507	9	AG346189 Mus muscu
C	28	228.4	15.4	1823	9	CU058635 CH216-88H
C	29	228	15.4	1259	9	CR004722 Forward s
C	30	228	15.4	2033	9	CU081977 CH216-165
C	31	228	15.4	2103	9	CU041093 CH216-51M
C	32	227.6	15.4	1712	9	CU078539 CH216-151
C	33	226.8	15.3	1547	9	CU066420 CH216-108
C	34	226.6	15.3	1377	9	CU077307 CH216-144
C	35	226.4	15.3	1716	8	CC222065 CH216-11A
C	36	225.6	15.2	1562	9	CU078533 CH216-151
C	37	225.6	15.2	1651	9	CU079030 CH216-154
C	38	225.4	15.2	1462	9	AG332292 Mus muscu
C	39	224.2	15.1	1506	9	CU082761 CH216-171
C	40	224	15.1	1536	9	CU078538 CH216-151
C	41	224	15.1	1628	9	CU082577 CH216-168
C	42	223.8	15.1	2001	9	CU118787 ISB1-72M2
C	43	223.6	15.1	1594	9	CU078613 CH216-151
C	44	223.6	15.1	1757	9	CU078645 CH216-152
C	45	223.4	15.1	1680	9	CU079033 CH216-154

#### ALIGNMENTS

RESULT 1  
LOCUS T09652 317 bp DNA linear GSS 05-DEC-2000  
DEFINITION 0224m3 gmbpH3.1, G. Roman Reddy Plasmodium falci-parum genomic  
ACCSSION T09652 GI:319484  
VERSION T09652.1 GI:319484  
KEYWORDS GSS.  
SOURCE Plasmodium falci-parum (malaria parasite P. falci-parum)  
ORGANISM Plasmodium falci-parum  
REFERENCE Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
AUTHORS 1 (bases 1 to 317)  
Reddy, G.R., Chakrabarti, D., Schuster, S.M., Ferl, R.J., Almira, E.C.  
and Dame, J.B.  
TITLE Gene sequence tags from Plasmodium falci-parum genomic DNA fragments  
JOURNAL prepared by the genease activity of mung bean nuclease  
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 90, 9867-9871 (1993)  
PUBMED 94052193  
COMMENT Other\_GSSs: 0224m7  
8234327  
Contact: Dame JB  
Dept. of Pathobiology, College of Veterinary Medicine  
University of Florida  
2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA  
Tel: 352 392 4700  
Fax: 352 392 9704  
Email: damejb@mail.vetmed.ufl.edu  
Seg primer: T3  
Class: shotgun.

#### FEATURES

source  
1..317  
/organism="Plasmodium falci-parum"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:5833"  
/clone="0224m"  
/lab\_host="E. coli XL1-Blue"  
/clone\_1lb="gmbpH3.1, G. Roman Reddy"  
/note="Vector: pBluescript SK(+); Genomic DNA, from  
asynchronous blood stage parasites of the cloned Honduran  
H3 isolate cultured in vitro, was digested at 500c  
nuclease in the presence of 30x formamide at 500c  
(Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.  
Nucleic Acids Research 16:6883-6896). The ends of the  
fragments were ligated to EcoR V-cleaved and  
dephosphorylated pBluescript SK(+). Recombinant plasmids  
transformed E. coli XL1-Blue."

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 14, 2005, 20:21:26 ; Search time 211.5 Seconds  
(without alignments)  
5027.299 Million cell updates/sec

Title: US-09-837-344-41

Sequence: 1 CAGAACCAACCAAGCATCT.....AATATTATTAACATATAA 1482

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MOBL=frame+ n2p.model -DEV=xlp  
-O=/cgn2.1/USPFO.spool.p/US09837344/runat.14012005.143841.4411/app.query.fasta\_1.1671  
-DBA=GeneSeq.23Sep04 -OPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=prco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09837344.CGN.1.1.236 @runat.14012005.143841.4411 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.GeneSeq.23Sep04:\*  
1: GeneSeqP1980s:\*  
2: GeneSeqP1990s:\*  
3: GeneSeqP2000s:\*  
4: GeneSeqP2001s:\*  
5: GeneSeqP2002s:\*  
6: GeneSeqP2003as:\*  
7: GeneSeqP2003bs:\*  
8: GeneSeqP2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2478	99.7	493	2	AAR26944
2	1640.5	66.0	456	8	ADO21942
3	1630	65.6	457	8	ADO21920
4	1561	62.8	316	2	AAR26941
5	969.5	39.0	462	2	AAR05766
6	770.5	31.0	318	2	AAR26943
7	603	24.3	117	4	AAR26937
8	376.5	14.3	419	4	ABG17301
9	376	15.1	1898	7	AAV30795
10	376	15.1	1898	7	ADD48869

11	361.5	14.5	412	2	AAW03626
12	359	14.4	1162	3	AAV96255
13	359	14.4	1162	3	AAV98500
14	359	14.4	1162	3	AAV98500
15	359	14.4	1162	5	ABE05621
16	359	14.4	1162	5	ABE05621
17	339.5	13.7	207	3	AAV44672
18	339.5	13.7	562	2	AAR70491
19	332.5	13.5	1192	3	AAI18165
20	331	13.3	2274	4	ABE58657
21	330.5	13.3	2990	8	ADP29884
22	327.5	13.2	360	2	AAW03627
23	323.5	13.0	611	2	AAV29039
24	323.5	13.0	611	4	AAU25510
25	323.5	13.0	611	7	ADG17213
26	323.5	13.0	611	7	ADG17213
27	320	12.9	658	4	ABE56532
28	320	12.9	800	4	ABE71459
29	320	12.9	3616	8	ADP45530
30	319	12.8	1972	2	AAW00024
31	318	12.8	554	4	ABE59454
32	317.5	12.8	1937	8	ADQ17289
33	317.5	12.8	1937	8	ADQ17289
34	316	12.7	1180	7	ADP06151
35	315.5	12.7	1170	5	ABE05710
36	315.5	12.7	1170	7	ADP06789
37	315.5	12.7	1170	7	ADP06789
38	315.5	12.7	1501	5	ABE59798
39	315.5	12.7	1564	6	ABP88812
40	315.5	12.7	1564	7	ADP06788
41	315	12.7	1960	4	AAW78854
42	315	12.7	1960	7	ADP70370
43	315	12.7	1960	7	ADP47946
44	315	12.7	1960	8	ADJ37149
45	315	12.7	2143	4	ABG01716

## ALIGNMENTS

RESULT 1	AAW03626	standard; protein; 493 AA.
ID	AAR26944	
AC	AAR26944;	
XX		
XX	25-MAR-2003 (revised)	
DT	08-FEB-1993 (first entry)	
XX		
DE	P. falciparum LSA gene C-terminal region.	
XX		
KW	Malaria; hepatocyte; sporozoite; T-cell epitope; B-cell epitope;	
KW	paludism; liver stage-specific antigen.	
XX		
OS	Plasmodium falciparum.	
XX		
FT	Key	Location/Qualifiers
FT	Region	13..213
FT	Region	/label= repetitive_region
FT	Region	214..493
FT	Region	/label= non-repetitive_region
XX		
XX	WO9213864-A1.	
XX		
PD	20-AUG-1992.	
XX		
XX	05-FEB-1992; 92WO-FR000104.	
PF		
XX	05-FEB-1991; 91FR-00001286.	
PR		
XX	(INSP ) INST PASTEUR.	
XX		
PA		
XX		
XX	Guerinmarchand C, Druilhe P;	
PI		
XX		

AAW03626 Human thy  
AAY96255 Kapozi's  
AAY58500 HIV8 ORF  
AAB62331 Amino aci  
ABD05621 Kapozi's  
ADJ65056 HIV8 late  
AAB44672 Gene 33 h  
AAR70491 Leucocyto  
AAB18165 Plasmodiu  
ABE58657 Drosophil  
ADP29884 Human bec  
AAW03627 Human fol  
AAY29039 T. gondii  
AAU25510 T. gondii  
ADG17213 T. gondii  
ABU53267 Human tes  
ABE56532 Drosophil  
ABD71459 Drosophil  
ADP45530 Chicken A  
AAW00024 Smooth mu  
ABE59454 Drosophil  
ADQ17289 Human sof  
ADQ17289 Human sof  
ADP06151 Bacterial  
ABE05710 Human tes  
ADP06789 Human pro  
ADP06789 Prostata  
ABP88812 Human pol  
ADP06788 Human str  
AAM78854 Human pro  
ADP70370 PAC 6802  
ADP47946 Human pro  
ADJ37149 Human mal  
ABG01716 Novel hum

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_nzp model

Run on: January 14, 2005, 20:38:17 ; Search time 44 Seconds

(without alignments)

4467,422 Million cell updates/sec

Title: US-09-837-344-41

Sequence: 1 CAGAGACACAAAGCATCT.....AATTTTATGAACATAA 1482

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 956278

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+nzp.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09837344/runat\_14012005\_143843\_4437/app\_query.fasta\_1.1671  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -LLIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09837344@cgn2\_1.1.46 -runat\_14012005\_143843\_4437 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:.\*  
5: /cgn2\_6/prodata/1/iaa/PCOTUS\_COMB.pep:.\*  
6: /cgn2\_6/prodata/1/iaa/backfill.pep:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1549	62.3	316	3	US-08-098-327E-31
2	1549	62.3	316	3	US-08-662-625-31
3	603	24.3	117	3	US-08-098-327E-20
4	603	24.3	117	3	US-08-662-625-20
5	554	22.3	107	3	US-08-098-327E-19
6	554	22.3	107	3	US-08-662-625-19
7	376	15.1	1898	2	US-08-056-200-94
8	376	15.1	1898	2	US-08-800-644-94
9	376	15.1	1898	4	US-09-538-092-1280
10	359	14.4	1162	2	US-08-728-323A-2
11	359	14.4	1162	2	US-09-298-568-2
12	359	14.4	1162	4	US-09-410-399-2

13	359	14.4	1162	4	US-09-894-273-2	Sequence 2, Appl1
14	336	13.5	608	4	US-09-270-767-32937	Sequence 32937, A
15	336	13.5	608	4	US-09-270-767-48154	Sequence 48154, A
16	323.5	13.0	611	4	US-09-216-393B-81	Sequence 81, Appl1
17	319	12.8	1972	4	US-08-875-435B-3	Sequence 3, Appl1
18	317.5	12.8	1937	4	US-09-538-092-918	Sequence 918, App
19	316	12.7	1180	4	US-09-543-681A-6336	Sequence 6436, Ap
20	315	12.7	1660	4	US-09-538-092-1077	Sequence 1077, Ap
21	314	12.6	1976	4	US-09-538-092-1078	Sequence 1078, Ap
22	309	12.4	1972	4	US-08-875-435B-4	Sequence 4, Appl1
23	305	12.3	1939	4	US-09-538-092-915	Sequence 915, App
24	303.5	12.2	924	4	US-09-248-796A-18798	Sequence 18798, A
25	303	12.2	1939	3	US-09-310-187A-1	Sequence 1, Appl1
26	303	12.2	1939	4	US-09-538-092-917	Sequence 917, App
27	301	12.1	1857	4	US-09-917-254-91	Sequence 91, Appl1
28	301	12.1	1972	4	US-09-538-092-1084	Sequence 1084, Ap
29	299.5	12.0	1886	3	US-08-938-105-3	Sequence 3, Appl1
30	299.5	12.0	1935	4	US-09-538-092-916	Sequence 916, App
31	299	12.0	3878	4	US-09-914-252-11	Sequence 11, Appl1
32	298	12.0	735	4	US-10-164-595-80	Sequence 80, Appl1
33	298	12.0	784	4	US-10-164-595-79	Sequence 79, Appl1
34	298	12.0	843	4	US-10-164-595-54	Sequence 54, Appl1
35	295	11.9	779	4	US-10-164-595-56	Sequence 56, Appl1
36	294	11.8	683	6	5210183-3	Patent No. 5210183
37	291	11.7	1239	4	US-09-688-188B-13	Sequence 13, Appl1
38	291	11.7	1239	4	US-09-291-417D-13	Sequence 13, Appl1
39	282.5	11.4	885	2	US-08-533-306A-4	Sequence 4, Appl1
40	282.5	11.4	885	2	US-08-742-923A-4	Sequence 4, Appl1
41	282	11.3	1581	4	US-09-866-108A-15754	Sequence 15754, A
42	282	11.3	2662	4	US-09-595-684B-31	Sequence 31, Appl1
43	280.5	11.3	2663	4	US-09-538-092-1252	Sequence 1252, Ap
44	279.5	11.2	630	4	US-09-248-796A-20275	Sequence 20275, A
45	278	11.2	1044	4	US-09-107-532A-5229	Sequence 5229, Ap

#### ALIGNMENTS

RESULT 1  
US-08-098-327E-31  
; Sequence 31, Application US/08098327E  
; Patent No. 6270771  
; GENERAL INFORMATION:  
; APPLICANT: GIERIN-MARCHAND, Claudine  
; APPLICANT: DRIITHE, Pierre  
; TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE  
; TITLE OF INVENTION: HEPATIC STAGES OF P. FALCIPARUM BEARING EPTIOPES CAPABLE  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/098,327E  
; FILING DATE: 24-NOV-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION NUMBER: FR 91 01286  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McGowan, Malcolm K.  
; REGISTRATION NUMBER: 39,300  
; REFERENCE/DOCKET NUMBER: 010830-045  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_nzp model

Run on: January 14, 2005, 21:00:12 ; Search time 211.5 Seconds

(without alignments)  
5063.176 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486  
Sequence: 1 CAAGACACAAAGCCATCT.....AATATTTTATGAACTATTA 1482

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 3216122

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+nzp.model -DEV=x1p  
-Q=/cgn2\_1/USPTO.spool/p/US09837344/runat\_14012005\_143844\_4457/app.query.fasta\_1.1671  
-DB=Published Applications AA -QPM=fastan , SUPPFX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62  
-TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09837344@CGN\_1\_1\_233-runat\_14012005\_143844\_4457  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOP=6 -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result

Query

No.	Score	Match	Length	DB	ID	Description
1	1549	62.3	316	9	US-09-837-344-31	Sequence 31, Appl
2	1549	62.3	316	10	US-09-900-963-31	Sequence 31, Appl
3	603	24.3	117	9	US-09-837-344-20	Sequence 20, Appl
4	603	24.3	117	10	US-09-900-963-20	Sequence 20, Appl
5	554	22.3	107	9	US-09-837-344-19	Sequence 19, Appl
6	554	22.3	107	10	US-09-900-963-19	Sequence 19, Appl
7	359	14.4	116	13	US-10-101-487-58	Sequence 58, Appl
8	359	14.4	116	11	US-09-894-273-2	Sequence 2, Appl
9	359	14.4	116	14	US-10-294-804-2	Sequence 2, Appl
10	359	14.4	116	17	US-10-194-046-2	Sequence 2, Appl
11	323.5	13.0	611	9	US-09-216-393-81	Sequence 81, Appl
12	323.5	13.0	611	14	US-10-321-856-81	Sequence 81, Appl
13	318.5	12.8	1961	14	US-10-028-248A-103	Sequence 103, Appl
14	318.5	12.8	1961	15	US-10-107-782-103	Sequence 103, Appl
15	317.5	12.8	1937	17	US-10-723-860-58	Sequence 58, Appl
16	317.5	12.8	1937	17	US-10-723-860-106	Sequence 106, Appl
17	315.5	12.7	1170	14	US-10-341-434-95	Sequence 95, Appl
18	315.5	12.7	1170	14	US-10-144-198-2	Sequence 2, Appl
19	315.5	12.7	1564	14	US-10-144-198-4	Sequence 4, Appl
20	315	12.7	1960	14	US-10-236-031B-62	Sequence 62, Appl
21	315	12.7	1960	14	US-10-028-248A-104	Sequence 104, Appl
22	315	12.7	1960	15	US-10-107-782-104	Sequence 104, Appl
23	313.5	12.6	1959	14	US-10-028-248A-16	Sequence 36, Appl
24	313.5	12.6	1959	15	US-10-107-782-36	Sequence 36, Appl
25	312.5	12.6	1959	14	US-10-028-248A-106	Sequence 106, Appl
26	312.5	12.6	1959	15	US-10-107-782-106	Sequence 106, Appl
27	311	12.5	240	13	US-10-101-487-75	Sequence 75, Appl
28	310.5	12.5	665	10	US-09-820-843A-107	Sequence 107, Appl
29	310	12.5	1963	14	US-10-369-493-5307	Sequence 5307, Appl
30	310	12.5	1963	14	US-10-369-493-5308	Sequence 5308, Appl
31	309.5	12.4	1956	14	US-10-369-493-6729	Sequence 6729, Appl
32	306.5	12.3	1790	14	US-10-369-493-1586	Sequence 1586, Appl
33	306	12.3	1999	14	US-10-028-248A-107	Sequence 107, Appl
34	306	12.3	1999	15	US-10-107-782-107	Sequence 107, Appl
35	305	12.3	621	15	US-10-108-260A-4409	Sequence 4409, Appl
36	305	12.3	1959	16	US-10-408-765A-2188	Sequence 2188, Appl
37	302.5	12.2	1130	16	US-10-369-493-6751	Sequence 6751, Appl
38	302.5	12.2	1959	17	US-10-798-037-4	Sequence 4, Appl
39	302	12.1	1958	16	US-10-408-765A-1629	Sequence 1629, Appl
40	301.5	12.1	1137	15	US-10-336-472-16	Sequence 16, Appl
41	301	12.1	1958	14	US-10-171-311-164	Sequence 164, Appl
42	301	12.1	1945	10	US-09-927-597-2	Sequence 2, Appl
43	301	12.1	1972	14	US-10-171-311-162	Sequence 162, Appl
44	301	12.1	1972	14	US-10-341-434-103	Sequence 103, Appl
45	301	12.1	1979	10	US-09-927-597-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-09-837-344-31  
; Sequence 31, Application US/09837344  
; Patent No. US20020041882A1  
GENERAL INFORMATION:  
APPLICANT: GUEBIN-MARCHAND, Claudine  
DRULHE, Pierre  
TITLE OF INVENTION: PEPTIDE SEQUENCES SPECIFIC FOR THE  
HEPTIC STAGES OF P. FALCIPARUM BEARING EPITOPES CAPABLE  
OF STIMULATING THE T LYMPHOCYTES

NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 14, 2005, 20:37:21 ; Search time 59.5 Seconds  
(without alignments)  
4793.050 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486  
Sequence: 1 CAAGAACACAAACGATCT.....AATATTATGAAACTATTA 1482

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xip  
-O=/cgn2/1/USPTO.spool.p/US09837344/runat.14012005.143842.4425/app.query.fasta\_1.1671  
-DB=PIR.79 -OPMT=fastan -SUFFIX=trp -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09837344 @CGN 1.1 80 @runat.14012005.143842.4425 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAIN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

PIR.79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2406.5	96.8	1909	2	A45592
2	380.5	15.3	1407	1	S28589
3	380.5	15.3	1549	1	A40691
4	376	15.1	1898	1	A45973
5	356	14.3	1017	2	T15598
6	353.5	14.2	737	2	T15597
7	332.5	13.4	1192	2	A71623
8	326.5	13.1	1979	1	S03166
9	323.5	13.0	1039	2	S18189
10	321	12.9	771	1	A33430
11	319	12.8	1938	2	UC5421
12	319	12.8	1972	2	UC5420
13	318.5	12.8	1961	1	A61231
14	317.5	12.8	1166	2	H86341

15	317.5	12.8	1937	2	I38055	myosin heavy chain
16	317.5	12.8	1992	2	A47297	myosin heavy chain
17	317.5	12.8	1432	2	B85431	trichohyalin like
18	316	12.7	1940	2	A29320	myosin heavy chain
19	314	12.6	1938	2	A59293	skeletal myosin he
20	314	12.6	1976	2	A59252	myosin heavy chain
21	313.5	12.6	1738	2	T14867	interaprin - slime
22	312.5	12.6	1959	1	A33977	myosin heavy chain
23	312	12.6	2007	1	B43402	myosin heavy chain
24	311.5	12.5	2139	2	T18296	myosin heavy chain
25	310.5	12.5	665	2	B71609	hypothetical prote
26	310.5	12.5	678	2	A54514	glutamic acid-rich
27	310	12.5	1963	1	W6M	myosin heavy chain
28	310	12.5	2116	2	A26655	myosin heavy chain
29	309.5	12.4	1956	2	T16416	hypothetical prote
30	309	12.4	1931	2	A59234	slow myosin heavy
31	309	12.4	1972	1	A41604	myosin heavy chain
32	307.5	12.4	1701	2	T09127	probable erythrocy
33	307	12.3	746	2	T47237	myosin II heavy ch
34	306.5	12.3	522	2	C96608	hypothetical prote
35	306.5	12.3	1790	2	S67593	transport protein
36	306	12.3	1999	1	S21801	myosin heavy chain
37	305.5	12.3	2017	1	A36014	myosin II heavy chain
38	305.5	12.3	2057	2	S61477	myosin II heavy ch
39	305	12.3	1964	2	A59282	nonmuscle myosin I
40	304.5	12.2	1939	1	A46762	myosin alpha heavy
41	304	12.2	451	2	G70241	hypothetical prote
42	303.5	12.2	1938	2	I49464	alpha cardiac myos
43	303.5	12.2	1939	2	I48175	myosin heavy chain
44	303	12.2	1536	2	A45605	mature-parasite-in
45	302.5	12.2	1130	2	T34081	hypothetical prote

# ALIGNMENTS

RESULT 1  
A45592  
Liver stage antigen LSA-1 - malaria parasite (Plasmodium falciparum)  
C:Species: Plasmodium falciparum  
C:Date: 22-Nov-1993 #sequence revision 02-Dec-1994 #text change 09-Jul-2004  
C:Accession: S24597; A45592; S29393; S34842; B45592; C45592; D45592  
R:Zhu, J., Hollingdale, M.  
Submitted to the EMBL Data Library, November 1990  
A:Reference number: S24597  
A:Accession: S24597  
A:Molecule type: DNA  
A:Residues: 1-1909 <ZHU>  
A:Cross-references: UNIPROT:Q25893; EMBL:X56203; NID:99915; PID:99916  
R:Zhu, J., Hollingdale, M.R., 223-226, 1991  
Mol. Biochem. Parasitol. 48, 223-226, 1991  
A:Title: Structure of Plasmodium falciparum liver stage antigen-1.  
A:Reference number: A45592; MUID:92107224; PMID:1840628  
A:Accession: A45592  
A:Molecule type: DNA  
A:Residues: 1-195;638-688;1165-1215;1590-1909 <ZH2>  
A:Note: Sequence extracted from NCBI Backbone (NCBIN:75010, NCBIN:75012, NCBIN:75014, R:Guerin-Marchand, C.; Drulhe, P.; Galey, B.; Londono, A.; Patapoutikou, J.; Beaudou Nature 329, 164-167, 1987  
A:Title: A liver-stage-specific antigen of Plasmodium falciparum characterized by gen  
A:Reference number: S29393; MUID:87315391; PMID:3306406  
A:Accession: S29393  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 333-387 <GUE1>  
A:Cross-references: EMBL:M28266  
R:Guerin-Marchand, C.; Drulhe, P.; Galey, B.; Londono, A.; Patapoutikou, J.; Beaudou submitted to the EMBL Data Library, April 1992  
A:Description: a liver-stage-specific antigen of Plasmodium falciparum characterized  
A:Reference number: S34842  
A:Accession: S34842  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 323-381, 'HKAI' <GUE2>

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: January 14, 2005, 20:22:46 ; Search time 306 Seconds

(without alignments)  
5573.237 Million cell updates/sec

Title: US-09-837-344-41

Perfect score: 2486

Sequence: 1 CAAGAACACACACGATCT.....AAATTTTATGAACTATAA 1482

Scoring table: BL0SUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp  
-Q/cgn2.1/usfpool.spool.p/us09837344/runat.14012005.143842.4417/app.query.fasta.1.1671  
-DB=UniProt\_02 -Qfmt=fstcan -SUPFIX=rup -TIMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-INITs-bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09837344@cgn2.1.1.307@runat.14012005.143842.4417 -NCPUS=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: UniProt\_02.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2478	99.7	493	2	Q25886 plasmodium
2	2406.5	96.8	1909	2	Q25893 plasmodium
3	2375.5	95.6	1596	2	Q81744 plasmodium
4	1451	58.4	280	2	Q25844 plasmodium
5	1450	58.3	280	2	Q25843 plasmodium
6	1447	58.2	280	2	Q25847 plasmodium
7	1447	58.2	280	2	Q25848 plasmodium
8	1447	58.2	280	2	Q25852 plasmodium
9	1445	58.1	280	2	Q25849 plasmodium
10	1445	58.0	280	2	Q25851 plasmodium
11	1443	58.0	280	2	Q25850 plasmodium
12	1442	57.8	280	2	Q25846 plasmodium
13	1437	57.7	280	2	Q25853 plasmodium
14	1434	57.7	280	2	Q25854 plasmodium
15	1433	57.6	280	2	Q25845 plasmodium
16	1432	57.6	280	2	Q25900 plasmodium

17	1430	57.5	280	2	Q25855 plasmodium
18	1430	57.5	280	2	Q25858 plasmodium
19	1428	57.4	280	2	Q27243 plasmodium
20	1426	57.4	280	2	Q25889 plasmodium
21	1366	54.9	264	2	Q9GTX5 plasmodium
22	770.5	31.0	318	2	Q25887 plasmodium
23	468.5	18.8	978	2	Q7BAK4 plasmodium
24	429.5	17.3	3427	2	Q61A79 plasmodium
25	429.5	17.3	3427	2	AA038039 plasmodium
26	398.5	16.0	1108	2	Q9ND10 plasmodium
27	393.5	15.8	1508	2	Q75JPS plasmodium
28	393.5	15.8	1508	2	AA545390 plasmodium
29	391.5	15.7	2349	2	Q81455 plasmodium
30	391	15.7	1750	2	Q8X0H2 plasmodium
31	389	15.6	826	2	Q9VPS3 plasmodium
32	381.5	15.3	1003	2	Q91LX9 plasmodium
33	381.5	15.3	1129	2	Q9GR71 plasmodium
34	380.5	15.3	1407	1	TRHY_RABIT
35	380.5	15.3	1549	1	TRHY_SHEEP
36	377.5	15.2	1089	2	Q40947 plasmodium
37	377	15.2	1036	2	Q9DUM3 plasmodium
38	376	15.1	1898	1	TRHY_HUMAN
39	370.5	14.9	10578	2	Q81SF5 plasmodium
40	370.5	14.9	18519	2	Q81SP6 plasmodium
41	370.5	14.9	18534	2	Q81SF7 plasmodium
42	367.5	14.8	1192	2	Q869E1 plasmodium
43	366	14.7	791	2	Q9DGL1 plasmodium
44	366	14.7	2760	2	Q815Y2 plasmodium
45	365	14.7	2275	2	Q81HVE plasmodium

## ALIGNMENTS

RESULT 1  
Q25886 PRELIMINARY: PRT: 493 AA.  
ID Q25886;  
AC Q25886;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Liver stage antigen-1 (Fragment).  
GN Name=LSA-1;  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94267224; PubMed=7515922;  
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,  
RA Melioux S., Guerin-Marchand C., Londono A., Raharimalala L.,  
RA Melioux S., Langsley G., Rousillon C., Tartar A., Druille P.,  
RA "The Plasmodium falciparum liver stage antigen-1 is well conserved and  
RT contains potent B and T cell determinants".  
RL J. Immunol. 153:190-204(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Fidock D.A., Gras-Masse H., Lepers J., Brahimi K., Benmohamed L.,  
RA Melioux S., Guerin-Marchand C., Londono A., Raharimalala L.,  
RA Melioux S., Langsley G., Rousillon C., Tartar A., Druille P.,  
RA "The Plasmodium falciparum liver stage antigen-1 is well conserved  
RT and harbors major B- and T-cell epitopes".  
RL J. Immunol. 0:0-0(1994).  
DR EMBL; Z30320; CAA82975.1; -.  
FT NON\_TER  
SQ SEQUENCE 493 AA; 59062 MW; FE252CC44F469965 CRC64;  
Alignment Scores:  
Pred. No.: 6.85e-107  
Score: 2478.00 Length: 493  
Percent Similarity: 100.00 Matches: 493  
Best Local Similarity: 100.00 Mismatches: 0  
Query Match: 99.68% Indels: 0  
Gaps: 0